All of Us RESEARCH PROGRAM | The Future of Health Begins With You

Updated April 27, 2018

Protocol v1 Summary

For the full protocol, please visit this link:
https://allofus.nih.gov/sites/default/files/aou_core_protocol_v1.7_mar_2018.pdf

What makes the *All of Us* Research Program unique?

- Size: 1 million or more participants
- Diversity: participants reflect the country's rich diversity, including communities historically underrepresented in biomedical research
- Scope: richness of information—clinical, biological, sociobehavioral, geographic, environmental, etc.
- Duration: longitudinal (lifelong): ongoing data/sample collection, in partnership with participants
- Accessibility: data available under agreed-upon terms to researchers from all sectors, including citizen scientists

	Protocol Component	Summary
Participants	**Primary Objective of the Study:** To build a robust research resource composed of participant-provided information (PPI), including environmental, physiologic, and health data and biospecimens from 1 million or more research participants who reflect the diversity of the U.S. This resource will facilitate the exploration of biological, social, and environmental determinants of health and disease.	
	Program Timeline and Enrollment	
	Program Timeline	• Program duration: 10 or more years. • Continuous updates through repeated participant engagement and sharing of participants' electronic health records (EHRs). • Biospecimen follow-up collection timeline is currently TBD.
	Participants	• Participants can enroll using a digital interface through two methods: as direct volunteers (DVs) or through participating health care provider organizations (HPOs).
	Getting to 1 Million	• Active enrollment through first five to six years of the program
	Inclusion/Exclusion Criteria	• Inclusion: adults older than 18 years with decisional capacity to consent and who currently reside within the U.S. or a U.S. territory. • Participation of children will begin within one year of national launch. • Exclusion: prisoners at time of enrollment.

Special Considerations	• Program will actively recruit diverse populations, including those that have been historically underrepresented in biomedical research. • Due to limitations in participant educational and consent materials, enrollment at initial program launch is restricted to participants who understand English or Spanish and can enroll digitally. • Over the first year of the program, more support for facilitated enrollment will be available for individuals who need it.
Informed Consent	
Methods of Consent	• e-Consent platform via program website and mobile apps, either virtually or at DV and HPO sites.
e-Consent Modules	• Two modules: consent to join the study and HIPAA/EHR authorization. Each module includes e-Consent screens, videos, electronic signature, and formative evaluation, all compliant with federal and state laws. • If participants do not wish to provide access to their EHR, they may say no. • If a participant says no to sharing their EHRs, they will likely not be invited to donate a biospecimen for genetic analyses and other assays.
Participant Data Collection	
Participant-Provided Information (PPI)	• Completed online or at DV/HPO site. • Questionnaire topics include sociodemographics, overall health, lifestyle, and health care access and utilization. More modules will be added over time.
Electronic Health Records	• At enrollment, participants will be asked to authorize sharing of their EHRs (if available) to *All of Us*. • EHR data will be sent from the participant's health care provider to the *All of Us* Data and Research Center (DRC). • Participants may also share health data with *All of Us* through Sync for Science or through data aggregators, such as PicnicHealth, as part of a pilot study. • Initial data types include all data available in the EHR, including demographics, visits, diagnoses, procedures, medications, laboratory visits, vital signs, and physician notes. They may include data about mental health, substance use, or HIV status.
Physical Measurements	• Participants who authorize EHR data sharing (regardless of whether they have current EHR data) may be invited to provide baseline physical measurements at their HPO or DV site. • Physical measurements taken include blood pressure, heart rate, height, weight, and waist and hip circumference. BMI is calculated automatically from measurements taken.

	Biospecimens	• Participants who authorize EHR data sharing may be invited to provide blood, urine, and/or saliva through their HPO enrollment center or designated DV biospecimen collection facility. The program also plans to develop the ability to do home visits when needed.
	Digital Health Data	• Additional mobile and digital health data from a yet-to-be-determined subset of participants may eventually be collected through health, wellness, and fitness devices; other sensors; and/or mobile applications.
	Patient Access to Data	• Participants will have access to data they provide directly, physical measurement data, biospecimen-related data, and genomic data (when available) through the participant portal (functionality to come). Physical measurement data is currently provided in print after completion of measurements.

Data	Protocol Component	Summary
	Data Collection, Accessibility, and Storage	
	Electronic Core Data Set	• A core data set of participant data, including PPI, physical measurements, biospecimen assays, and EHR information (when available), will be developed and available to researchers. • Obvious identifiers will be removed from the data that researchers access. • The core data set will be data linked by geolocation, based on participant-provided residential and work addresses. • Additional potential sources of data links to the core data set for future inclusion may include Social Security Death Master Files, pharmacy system data, claims data, and health registry data.
	Accessing Core Data Set	• Academic scientists, commercial organizations, and interested citizen scientists will need to request access to the data. The Research Portal will open a year to a year and a half after program launch. • Approved users will be able to query the data and run analyses using the cloud infrastructure. • Data will be made available using a researcher-based (not study-based) mechanism. Tier-specific data passports granted to approved users will allow access to three different tiers of data: public access data, registered access data, and controlled access data.
	Biospecimen Collection and Storage All biospecimen samples are initially processed at the site of collection, transported at 4ºC, and stored at –80ºC at the Mayo Clinic Biobank.	

Blood	• About 50 mL per visit, collected in Clot Activator (SST), Plasma Separator (PST), EDTA, circulating cell-free DNA, and PAXgene tubes. • If a participant has donated blood within the past week or received a blood transfusion within the past six months, they will be asked to reschedule the collection.
Urine	• About 50 mL collected per visit.
Saliva	• Saliva may be collected as an alternative to a blood sample if the participant is unable to return for a second collection visit following an unsuccessful initial collection visit or if the blood sample collection is unsuccessful after two attempts.
Data from Biospecimens	• Genomic data will be generated for those who consent, with a pilot on the responsible return of genetic results once genomic data are generated. • Decisions about other data types generated from biospecimens are yet to be determined. The recent IdeaScale campaign and Research Priorities Workshop in March 2018 are the first steps in gathering stakeholder input.
Accessing Data	• Approved users will be able to access the generated data by a short application process that will include ethics training, agreement to a code of conduct, and, in some cases, sign-off from their institution. • Data will be accessed and analyzed through the *All of Us* research environment.
Accessing Biospecimens	• Requests for access will be submitted to *All of Us*. The policies and processes for this are being developed.
Data Security and Protection	• Participant data elements will be transferred through encrypted channels to the core data set, which will be stored at the DRC. The data will be stored in a secured cloud-computing environment that follows rigorous standards to protect individual privacy and data confidentiality. Data cannot be downloaded.